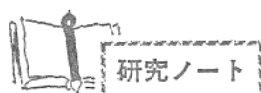


モデリングの方法論に関する研究



田村 坦 之*

「モデリング」とか「モデル化」という用語が、我々「システム」に携わる者の間では日常茶飯に使われている。工学的なシステムはもとより、例えば社会的なシステムに関連してシステム分析を行う際に、システムの「モデル」が果たす役割の重要性には計り知れないものがある。モデルとは「実システムを単純化し、様式化して表現したもので、対象とする問題の本質的な因果関係を引き出したもの」であり、例えば一連の数式、コンピュータ・プログラム、グラフ、シナリオ（言葉のみによる叙述）など様々なものがある。

制御工学の分野では、制御対象の動特性を知る上で「システム同定」¹⁾に関する研究が数多く行われている。そこでは、システム同定とは「実システムの入出力データをもとにして、与えられた数式モデルのクラスから実システムと等価と見做せる数式モデルを見つけること」と定義されており²⁾、これは制御対象のモデリングを意味している。ただし、制御工学者は制御対象に関する個別分野の理論すなわち現象論とは無関係に、システムの入出力データの解析のみに頼ってモデルを導入しようとする傾向が強い。この結果、個別分野を理論的に研究している人達からは「理論なきモデル」としてひんしゆくを買う結果に陥るので注意を要する。

一方、実システムにおいて、各個別分野の理論的な解析のみによってモデリングが完結する場合は稀である。まず理論的解析によってモデルの構造を構築し、モデルの次数や未知パラメータの値を入出力データの解析によって求めるといった具合に、個別分野の理論とシステム的方法論とが相補的に相手の弱点を補い合うこ

とによって有用なモデリングが可能となる場合が多い³⁾⁴⁾。

以下、特に対象を限定せずに、モデリングの方法論に関する最近のトピックスと、それに関連して最近筆者らの所で行った研究の一端を紹介する。

1. 静的モデリング：重回帰分析における変数選択と改良形 GMDH

一般に、従属形の静的モデルとして、出力変数（従属変数、目的変数） y と入力変数（独立変数、説明変数） (x_1, x_2, \dots, x_p) に対する数式モデル

$$y = f(x_1, x_2, \dots, x_p) + \varepsilon \quad \dots \dots (1)$$

を仮定し、 n 組のデータ $\{x_{\alpha 1}, x_{\alpha 2}, \dots, x_{\alpha p}; y_{\alpha}\}$ ($\alpha=1, 2, \dots, n$) を用いてモデリングを行う。特に(1)式が未知パラメータに関して線形であると仮定できる場合には、重回帰分析によって重回帰モデル(1)式に対するあてはめを行うことができる。しかし、入力変数の個数 p が20以上にもなる場合には、この入力変数すべてに対して重回帰分析を行うことは労多くして効果が少なく、モデルが複雑になり過ぎてモデルの信頼性を低下させる結果になる。そこで、 p 個の入力変数自体どのようなものを採用すべきか、すなわち、重回帰分析に先立って変数選択の問題を解決しておくことがモデリングの立場から極めて重要となる。⁵⁾

対象としている分野の現象論的な考察からある程度の変数選択は可能であるが、さらに、予測平方和 (PSS)、自乗平均予測誤差 (MSEP)、赤池氏の情報量規準 AIC⁶⁾、BIC⁷⁾などを評価基準として変数選択を行う方法が考えられている。

Ivakhnenko が提案した GMDH (Group Method of Data Handling)⁸⁾は、(1)式として Kolmogorov-Gabor の多項式（これを完全表

* 田村坦之 (Hiroyuki TAMURA), 大阪大学工学部, 精密工学科, 助教授, 工学博士, 応用システム解析。

現式と呼ぶ.)

$$y = a_0 + \sum_{i=1}^p a_i x_i + \sum_{i=1}^p \sum_{j=1}^p a_{ij} x_i x_j + \dots \quad (2)$$

を仮定し、2つの入力変数 x_i, x_j の2次多項式(これを部分表現式と呼ぶ.)

$$z_k = b_0 + b_1 x_i + b_2 x_j + b_3 x_i x_j + b_4 x_i^2 + b_5 x_j^2 \dots \quad (3)$$

$$k=1, 2, \dots, p(p-1)/2$$

を多層に積重ねてモデリングを行う方法である。そこでは入出力データの一部(trainingデータ)を用いて部分表現式のパラメータ推定を行い、残りのデータ(checkingデータ)を用いて中間変数 z_k の選択と多層構造の層の打ち切りとを行っている。この結果、重回帰分析にくらべて、モデリングに必要なデータの個数はずっと少なく済み、多変数で非線形のモデリングが容易に行えるという利点を得ている。その反面、

- (1) 同定されるモデルは不必要に高次の多項式になり、入力変数の個数が多くて次数の低いモデルは出てこない。
- (2) 入出力データの分割の仕方によって、得られる結果が大巾に異なる。
- (3) 各層で中間変数を幾つ選択するかによって、出てくる結果が大巾に異なる。

といった欠点を含んでいる。筆者らの所では、これらの問題点を解決する改良形 GMDH⁹⁾¹⁰⁾として「入出力データを分割せずすべてのデータを用いてモデルの予測誤差を推定し、この推定値を最小にするような部分表現式の変数選択、中間変数の選択、多層構造の層の打ち切りを行う方法」を提案している⁹⁾¹⁰⁾。この方法論の有効性は、(i)簡単な力学系のモデリング⁹⁾、(ii)広域大気汚染濃度パターンのモデリング¹⁰⁾、などへの応用例を通じて検証している。

2. 動的モデリング：時系列解析と逐次形赤池法

動的モデリングの方法論には、大別して時系列解析と、システムの実現問題として捕える方法とがある。時系列解析とは、動的システムの入出力データのみに着目して、自己回帰(AR)モデル、移動平均(MA)モデル、自己回帰移動平均(ARMA)モデル、自己回帰集積移動

平均(ARIMA)モデルなどによってモデリングを行う方法論を意味している。ARモデルによる赤池法¹¹⁾、ARIMAモデルによるBox-Jenkins法¹²⁾などが有名であるが、中でも最終予測誤差(FPE)を用いてモデルの次数を自己選択する赤池法は実用性が高く、多方面に応用されている。時系列解析におけるモデル次数の選択は、重回帰分析における変数選択に対応しており、FPEはARモデルに対するAICを表している。

赤池法は、一括して与えられた時系列データをもとにしてモデリングを行う方法を与えているが、オンライン・リアルタイムでモデリングを行う実際的な場面を想定すると、新たに一つあるいは複数個のデータが時々刻々得られるたびに逐次的にモデルを更新する方法論が欲しくなる。この目的にそって筆者らの所で開発した手法が逐次形赤池法である¹³⁾。この方法は、データの数が非常に多くなる場合や、データ収集に時間がかかる場合に特に有効である。逐次形赤池法の有効性は、河川水質システムのモデリングへの応用によって検証している⁴⁾¹⁴⁾。

3. 構造的モデリング：ISMとIPSM

複雑な社会システムや経営・管理システムを対象にして、システムを構成している要素間の因果関係に関する構造を有向グラフの形でモデル化する方法論が注目をあびている。Battelle研究所で開発されたISM(Interpretive Structural Modeling)¹⁵⁾やDEMATEL(Decision Making Trial and Evaluation Laboratory)がその一例である。

最近、筆者らの所では「或る目的を達成するための幾つかの代替案の中から、意思決定者(以下DMと略す)の選好に合致した好ましい行動案を選択するシステム的方法論」の一つとしてIPSM(Interpretive Preference Structural Modeling)¹⁶⁾という方法論を提案している。IPSMはISMの手法を応用して、複数の代替案に対するDMの選好構造のモデリングを行うことを目的としている。まず、 p 個の代替案 A_1, A_2, \dots, A_p (多目的システムでは p 個のパレート最適解)に対するDMの対比較によって $p \times p$ の選好構造行列 $M=[m_{ij}]$ を

作る。ただし、

(i) 代替案 A_i が A_j より好ましいとき

$$m_{ij}=0, m_{ji}=1$$

(ii) 代替案 A_j が A_i より好ましいとき

$$m_{ij}=1, m_{ji}=0$$

(iii) 代替案 A_i と A_j の好ましさが同等のとき

$$m_{ij}=1, m_{ji}=1$$

(iv) 代替案 A_i と A_j の好ましさが比較できないとき

$$m_{ij}=0, m_{ji}=0$$

とする。代替案の数が多いときには、すべての m_{ij} の値をDMにたずねるのは大変であるが、選好関係に関する推移律を仮定するならば、DMとコンピュータとの対話によってDMに対する質問回数を大巾に削減することができる。選好構造行列Mをグラフ理論によって解析し、代替案の間の選好関係を階層的な有向グラフとして図示する。ただ、この有向グラフにはDM自ら直接判断した選好関係と、推移律によって間接的に推定した選好関係の両方を含んでいるので、DMとコンピュータとの対話によって逐次修正するという一種の学習過程を繰り返した末、最終的にDMの実質的な選好構造に収束する。

ここでDMとは、単に政策決定者だけを意味するのではなく、最近新聞紙上を賑わしている環境アセスメントを例にとると「住民」がDMの役割を果すことも有り得る。このような場合には地域住民に対するアンケート調査や公聴会によって選好構造行列Mを作成することが考えられる。

以上、浅学非才を顧みず、モデリングの方法論に関する最近の研究のごく一端を概観した。各分野の専門家が分野別にモデリングを議論しているすぐれた文献に、「計測と制御、15巻、3号、昭. 51 (モデリングとシミュレーション特集号)」がある。モデリングの方法論については、まだまとまったテキストも世に出ていない状態で、今後、方法論の整理・統合が望まれる。おわりに、本研究中、環境問題に関連して筆者らの所で行った研究は、文部省科研特定研究(1) (代表者榎木義一教授)に関連して行った

研究であることを付記し謝意を表する。

参考文献

- 1) P. Eykhoff: *System Identification*, John-Wiley, 1974.
- 2) L. A. Zadeh: From circuit theory to system theory, *Proc. IRE*, **50**, 856/865, 1962.
- 3) S. Rinaldi, R. Soncini-Sessa, H. Stehfest and H. Tamura: *Modeling and Control of River Quality*, McGraw-Hill, 1978 (forthcoming).
- 4) H. Tamura: On modeling and identification of river quality systems using distributed-lag models, *IFIP Working Conf. on Modeling and Simulation of Land, Air and Water Resources Systems*, Ghent, Belgium, Aug. 30—Sept. 2, 1977.
- 5) 奥野忠一他: 続多変量解析法, 日科技連, 昭. 51.
- 6) 赤池弘次: 情報量規準 AIC とは何か, 数理科学, 153号, 5/11, 昭. 51.
- 7) H. Akaike: On newer statistical approaches to parameter estimation and structure determination (invited paper), *7-th IFAC World Congress*, Helsinki, June 1979. (forthcoming).
- 8) 池田・榎木: GMDH と複雑な系の同定・予測, 計測と制御, 14巻, 2号, 185/195, 昭. 50.
- 9) 近藤・田村: 情報量規準 AIC を用いた改良形 GMDH, 第3回システムズ・シンポジウム前刷 (計測自動制御学会主催), 163/168, 昭. 52-7.
- 10) H. Tamura and T. Kondo: Large-spatial pattern identification of air pollution by a combined model of source-receptor matrix and revised GMDH, *Proc. IFAC Symp. on Environmental Systems Planning, Design and Control*, 373/380, Kyoto, Aug. 1977.
- 11) 赤池・中川: ダイナミックシステムの統計的解析と制御, サイエンス社, 昭. 47.
- 12) G. E. P. Box and G. M. Jenkins: *Time Series Analysis—Forecasting and Control*, Holden-Day, 1970.
- 13) 田村・川口: 多次元自己回帰モデルによるオンライン逐次形システム同定—赤池法のオンライン化—, 計測自動制御学会論文集, 13巻, 1号, 14/20, 昭. 52.
- 14) H. Tamura and T. Kawaguchi: Real-time estimation of orders and parameters of distributed-lag models for river quality, in A. Szöllösi-Nagy, et al ed. *Recent Developments in Real-Time Forecasting/Control of Water Resources Systems*, (Proc. IIASA/WMO Workshop), John-Wiley, 1978 (forthcoming).
- 15) 河村和彦: 複雑な社会問題を取扱う一手法: Interpretive Structural Modeling, 計測と制御, 16巻, 1号, 157/161, 昭. 52.
- 16) 田村・榎井: 多目的システムにおけるトレード・オフ分析—ISMによる選好構造把握—, 文部省科研総合研究(A)動学的経済システムの決定と制御 (代表者妙見孟教授), 研究報告書, 昭. 52-2.